Uniform mixing and completely positive sofic entropy

Tim Austin and Peter Burton

September 20, 2018

Abstract

Let G be a countable discrete sofic group. We define a concept of uniform mixing for measure-preserving G-actions and show that it implies completely positive sofic entropy. When G contains an element of infinite order, we use this to produce an uncountable family of pairwise nonisomorphic G-actions with completely positive sofic entropy. None of our examples is a factor of a Bernoulli shift.

1 Introduction

Let G be a countable discrete sofic group, (X, μ) a standard probability space and $T : G \curvearrowright X$ a measurable G-action preserving μ. In [2], Lewis Bowen defined the sofic entropy of (X, μ, T) relative to a sofic approximation under the hypothesis that the action admits a finite generating partition. The definition was extended to general (X, μ, T) by Kerr and Li in [9] and Kerr gave a more elementary approach in [8]. In [3] Bowen showed that when G is amenable, sofic entropy relative to any sofic approximation agrees with the standard Kolmogorov-Sinai entropy. Despite some notable successes such as the proof in [2] that Bernoulli shifts with distinct base-entropies are nonisomorphic, many aspects of the theory of sofic entropy are still relatively undeveloped.

Rather than work with abstract measure-preserving G-actions, we will use the formalism of G-processes. If G is a countable group and A is a standard Borel space, we will endow A^G with the right-shift action given by $(g \cdot a)(h) = a(hg)$ for $g, h \in G$ and $a \in A^G$. A G-process over A is a Borel probability measure μ on A^G which is invariant under this action. Any measure-preserving action of G on a standard probability space is measure-theoretically isomorphic to a G-process over some standard Borel space A. We will assume the state space A is finite, which corresponds to the case of measure-preserving actions which admit a finite generating partition. Note that by results of Seward from [12] and [13], the last condition is equivalent to an action admitting a countable generating partition with finite Shannon entropy.

In [1], the first author introduced a modified invariant called model-measure sofic entropy which is a lower bound for Bowen's sofic entropy. Let $\Sigma = (\sigma_n : G \to \mathrm{Sym}(V_n))$ be a sofic approximation to G. Model-measure sofic entropy is constructed in terms of sequences $(\mu_n)_{n=1}^{\infty}$ where μ_n is a probability measure on A^{V_n}. If these measures replicate the process (A^G, μ) in an appropriate sense then we say that $(\mu_n)_{n=1}^{\infty}$ locally and empirically converges to μ. We refer the reader to [1] for the precise definitions. We have substituted the phrase 'local and empirical convergence' for the phrase 'quenched convergence' which appeared in [1]. This has been done to avoid confusion with an alternative use of the word 'quenched' in the physics literature. A process is said to have completely positive model-measure sofic entropy if every nontrivial factor has positive model-measure sofic entropy. The goal of this paper is the to prove the following theorem, which generalizes the main theorem of [5].

Theorem 1.1. *Let G be a countable sofic group containing an element of infinite order. Then there exists an uncountable family of pairwise nonisomorphic G-processes each of which has completely positive model-measure sofic entropy (and hence completely positive sofic entropy) with respect to any sofic approximation to G. None of these processes is a factor of a Bernoulli shift.*

In order to prove Theorem 1.1 we introduce a concept of uniform mixing for sequences of model-measures. This uniform model-mixing will be defined formally in Section 3. It implies completely positive model-measure sofic entropy.

Theorem 1.2. *Let G be a countable sofic group and let (A^G, μ) be a G-process with finite state space A. Suppose that for some sofic approximation Σ to G, there is a uniformly model-mixing sequence $(\mu_n)_{n=1}^{\infty}$ which locally and empirically converges to μ over Σ. Then (A^G, μ) has completely positive lower model-measure sofic entropy with respect to Σ.*

As in [5], the examples we exhibit to establish Theorem 1.1 are produced via a coinduction method for lifting H-processes to G-processes when $H \leq G$. If (A^H, ν) is an H-process then we can construct a corresponding G-process (A^G, μ) as follows. Let T be a transversal for the right cosets of H in G. Identify G as a set with $H \times T$ and thereby identify A^G with $(A^H)^T$. Set $\mu = \nu^T$. We call (A^G, μ) the coinduced process and denote it by $\mathrm{CInd}_H^G(\nu)$. (See page 72 of [7] for more details on this construction.) When $H \cong \mathbb{Z}$ this procedure preserves uniform mixing.

Theorem 1.3. *Let G be a countable sofic group and let $(A^{\mathbb{Z}}, \nu)$ be a uniformly mixing $\mathbb{Z}$-process with finite state space A. Let $H \leq G$ be a subgroup isomorphic to $\mathbb{Z}$ and identify $A^{\mathbb{Z}}$ with A^H. Then for any sofic approximation Σ to G, there is a uniformly model-mixing sequence of measures which locally and empirically converges to $\mathrm{CInd}_H^G(\nu)$ over Σ.*

We remark that it is easy to see that if (A^G, μ) is a Bernoulli shift (that is to say, μ is a product measure), then there is a uniformly model-mixing sequence which locally and empirically converges to μ. Indeed, if $\mu = \eta^G$ for a measure η on A then the measures η^{V_n} on A^{V_n} are uniformly model-mixing and locally and empirically converge to μ. Thus Theorem 1.2 shows that Bernoulli shifts with finite state space have completely positive sofic entropy, giving another proof of this case of the main theorem from [10]. We believe that completely positive sofic entropy for general Bernoulli shifts can be deduced along the same lines, requiring only a few additional estimates, but do not pursue the details here.

1.1 Acknowledgements

The first author's research was partially supported by the Simons Collaboration on Algorithms and Geometry. The second author's research was partially supported by NSF grants DMS-0968710 and DMS-1464475.

2 Preliminaries

2.1 Notation

The notation we use closely follows that in [1]; we refer the reader to that reference for further discussion. Let A be a finite set. For any pair of sets $W \subseteq S$ we let $\pi_W : A^S \to A^W$ be projection onto the W-coordinates (thus our notation leaves the larger set S implicit). Let G be a countable group and let (A^G, μ) be a G-process. For $F \subseteq G$ we will write $\mu_F = \pi_{F*}\mu \in \mathrm{Prob}(A^F)$ for the F-marginal of μ.

Let B be another finite set and let $\phi : A^G \to B$ be a measurable function. If $F \subseteq G$ we will say that ϕ is F-local if it factors through π_F. We will say ϕ is local if it is F-local for some finite F. Let $\phi^G : A^G \to B^G$ be given by $\phi^G(a)(g) = \phi(g \cdot a)$ and note that ϕ^G is equivariant between the right-shift on A^G and the right-shift on B^G.

Let V be a finite set and let σ be a map from G to $\mathrm{Sym}(V)$. For $g \in G$ and $v \in V$ we write $\sigma^g \cdot v$ instead of $\sigma(g)(v)$. For $F \subseteq G$ and $S \subseteq V$ we define

$$\sigma^F(S) = \{\sigma^g \cdot s : g \in F, s \in S\}.$$

For $v \in V$ we write $\sigma^F(v)$ for $\sigma^F(\{v\})$. We write $\Pi^\sigma_{v,F}$ for the map from A^V to A^F given by $\Pi^\sigma_{v,F}(\overline{a})(g) = \overline{a}(\sigma^g \cdot v)$ for $\overline{a} \in A^V$ and $g \in F$. We write Π^σ_v for $\Pi^\sigma_{v,G}$. With $\phi : A^G \to B$ as before, we write ϕ^σ for the map from A^V to B^V given by $\phi^\sigma(\overline{a})(v) = \phi\big(\Pi^\sigma_v(\overline{a})\big)$.

If D is a finite set and η is a probability measure on D then $\mathrm{H}(\eta)$ denotes the Shannon entropy of η, and for $\epsilon > 0$ we define

$$\mathrm{cov}_\epsilon(\eta) = \min\big\{|F| : F \subseteq D \text{ is such that } \eta(F) > 1 - \epsilon\big\}.$$

If $\phi : D \to E$ is a map to another finite set then we may write $\mathrm{H}_\mu(\phi)$ in place of $\mathrm{H}(\phi_*\mu)$. For $p \in [0,1]$ we let $\mathrm{H}(p) = -p \log p - (1-p)\log(1-p)$.

We use the $o(\cdot)$ and $\lesssim$ asymptotic notations with respect to the limit $n \to \infty$. Given two functions f and g on $\mathbb{N}$, the notation $f \lesssim g$ means that there is a positive constant C such that $f(n) \leq Cg(n)$ for all n.

2.2 An information theoretic estimate

Lemma 2.1. *Let A be a finite set and let $(V_n)_{n=1}^\infty$ be a sequence of finite sets such that $|V_n|$ increases to infinity. Let μ_n be a probability measure on A^{V_n}. We have*

$$\liminf_{n\to\infty} \frac{\mathrm{H}(\mu_n)}{|V_n|} \leq \sup_{\epsilon > 0} \liminf_{n\to\infty} \frac{1}{|V_n|} \log \mathrm{cov}_\epsilon(\mu_n).$$

Proof. Let μ be a probability measure on a finite set F and let $E \subseteq F$. By conditioning on the partition $\{E, F \setminus E\}$ and then recalling that entropy is maximized by uniform distributions we obtain

$$\begin{aligned} \mathrm{H}(\mu) &= \mu(E) \cdot \mathrm{H}(\mu(\cdot\,|\,E)) + \mu(F \setminus E) \cdot \mathrm{H}(\mu(\cdot\,|\,F \setminus E)) + \mathrm{H}(\mu(E)) \\ &\leq \mu(E) \cdot \log(|E|) + (1 - \mu(E)) \cdot \log(|F \setminus E|) + \mathrm{H}(\mu(E)). \end{aligned} \tag{2.1}$$

Now let μ_n and V_n be as in the statement of the lemma. Let $\epsilon > 0$ and let $S_n \subseteq A^{V_n}$ be a sequence of sets with $\mu_n(S_n) > 1-\epsilon$ and $|S_n| = \mathrm{cov}_\epsilon(\mu_n)$. By applying (2.1) with $F = A^{V_n}$ and $E = S_n$ we have

$$\begin{aligned}
\liminf_{n\to\infty} \frac{\mathrm{H}(\mu_n)}{|V_n|} &\le \liminf_{n\to\infty} \frac{1}{|V_n|}\big(\mu(S_n)\cdot\log(|S_n|) + (1-\mu(S_n))\cdot\log(|A^{V_n}\setminus S_n|) + \mathrm{H}(\mu(S_n))\big)\\
&\le \liminf_{n\to\infty} \frac{1}{|V_n|}\big(\log(|S_n|) + \epsilon\cdot\log(|A^{V_n}|) + \mathrm{H}(\epsilon)\big)\\
&\le \left(\liminf_{n\to\infty} \frac{1}{|V_n|}\log\mathrm{cov}_\epsilon\big(\mu_n\big)\right) + \epsilon\cdot\log(|A|).
\end{aligned}$$

Now let ϵ tend to zero to obtain the lemma. $\square$

3 Metrics on sofic approximations and uniform model-mixing

Let us fix a proper right-invariant metric ρ on G: for instance, if G is finitely generated then ρ can be a word metric, and more generally we may let $w : G \to [0,\infty)$ be any proper weight function and define ρ to be the resulting weighted word metric. Again let V be a finite set and let σ be a map from G to $\mathrm{Sym}(V)$. Let H_σ be the graph on V with an edge from v to w if and only if $\sigma^g\cdot v = w$ or $\sigma^g \cdot w = v$ for some $g\in G$. Define a weight function W on the edges of H_σ by setting

$$W(v,w) = \min\big\{\rho(g,1_G) : \sigma^g\cdot v = w \text{ or } \sigma^g\cdot w = v\big\}.$$

If v and w are in the same connected component of H_σ let ρ_σ be the W-weighted graph distance between v and w, that is

$$\rho_\sigma(v,w) = \min\left\{\sum_{i=0}^{k-1} W(p_i,p_{i+1}) : (v = p_0, p_1, \ldots, p_{k-1}, p_k = w) \text{ is an } H_\sigma\text{-path from } v \text{ to } w\right\}.$$

Having defined ρ_σ on the connected components of H_σ, choose some number M much larger than the ρ_σ-distance between any two points in the same connected component. Set $\rho_\sigma(v,w) = M$ for any pair v,w of vertices in distinct connected components of H_σ. Note that if $(\sigma_n : G\to\mathrm{Sym}(V_n))$ is a sofic approximation to G then for any fixed $r<\infty$ once n is large enough the map $g\mapsto\sigma_n^g\cdot v$ restricts to an isometry from $B_\rho(1_G,r)$ to $B_{\rho_{\sigma_n}}(v,r)$ for most $v\in V_n$.

In the sequel the sofic approximation will be fixed, and we will abbreviate ρ_{σ_n} to ρ_n. We can now state the main definition of this paper.

Definition 3.1. *Let $(V_n)_{n=1}^\infty$ be a sequence of finite sets with $|V_n|\to\infty$ and for each n let σ_n be a map from G to $\mathrm{Sym}(V_n)$. Let A be a finite set. For each $n\in\mathbb{N}$ let μ_n be a probability measure on A^{V_n}. We say the sequence $(\mu_n)_{n=1}^\infty$ is* ***uniformly model-mixing*** *if the following holds. For every finite $F\subseteq G$ and every $\epsilon>0$ there is some $r<\infty$ and a sequence of subsets $W_n\subseteq V_n$ such that*

$$|W_n| = (1-o(1))|V_n|$$

and if $S\subseteq W_n$ is r-separated according the metric ρ_n then

$$\mathrm{H}\left(\pi_{\sigma_n^F(S)*}\mu_n\right) \ge |S|\cdot(\mathrm{H}(\mu_F)-\epsilon).$$

This definition is motivated by Weiss' notion of uniform mixing from the special case when G is amenable: see [14] and also Section 4 of [5]. Let us quickly recall that notion in the setting of a G-process (A^G, μ). First, if $K \subseteq G$ is finite and $S \subseteq G$ is another subset, then S is K**-spread** if any distinct elements $s_1, s_2 \in S$ satisfy $s_1 s_2^{-1} \notin K$. The process (A^G, μ) is **uniformly mixing** if, for any finite-valued measurable function $\phi : A^G \to B$ and any $\epsilon > 0$, there exists a finite subset $K \subseteq G$ with the following property: if $S \subseteq G$ is another finite subset which is K-spread, then

$$\mathrm{H}\big((\phi^G_* \mu)_S\big) \geq |S| \cdot (\mathrm{H}_\mu(\phi) - \epsilon).$$

Beware that we have reversed the order of multiplying s_1 and s_2^{-1} in the definition of 'K-spread' compared with [5]. This is because we work in terms of observables such as ϕ rather than finite partitions of A^G, and shifting an observable by the action of $g \in G$ corresponds to shifting the partition that it generates by g^{-1}.

The principal result of [11] is that completely positive entropy implies uniform mixing. The reverse implication also holds: see [6] or Theorem 4.2 in [5]. Thus, uniform mixing is an equivalent characterization of completely positive entropy.

The definition of uniform mixing may be rephrased in terms of our proper metric ρ on G as follows. The process (A^G, μ) is uniformly mixing if and only if, for any finite-valued measurable function $\phi : A^G \to B$ and any $\epsilon > 0$, there exists an $r < \infty$ with the following property: if $S \subseteq G$ is r-separated according to ρ, then

$$\mathrm{H}\big((\phi^G_* \mu)_S\big) \geq |S| \cdot (\mathrm{H}_\mu(\phi) - \epsilon).$$

This is equivalent to the previous definition because a subset $S \subseteq G$ is r-separated according to ρ if and only if it is $B_\rho(1_G, r)$-spread. The balls $B_\rho(1_G, r)$ are finite, because ρ is proper, and any other finite subset $K \subseteq G$ is contained in $B_\rho(1_G, r)$ for all sufficiently large r.

This is the point of view on uniform mixing which motivates Definition 3.1. We use the right-invariant metric ρ rather than the general definition of 'K-spread' sets because it is more convenient later.

Definition 3.1 is directly compatible with uniform mixing in the following sense. If G is amenable and $(F_n)_{n=1}^\infty$ is a Følner sequence for G, then the sets F_n may be regarded as a sofic approximation to G: an element $g \in G$ acts on F_n by translation wherever this stays inside F_n and arbitrarily at points which are too close to the boundary of F_n. If (A^G, μ) is an ergodic G-process, then it follows easily that the sequence of marginals μ_{F_n} locally and empirically converge to μ over this Følner-set sofic approximation. If (A^G, μ) is uniformly mixing, then this sequence of marginals is clearly uniformly model-mixing in the sense of Definition 3.1.

On the other hand, suppose that (A^G, μ) admits a sofic approximation and a locally and empirically convergent sequence of measures over that sofic approximation which is uniformly model-mixing. Then our Theorem 1.2 shows that (A^G, μ) has completely positive sofic entropy. If G is amenable then sofic entropy always agrees with Kolmogorov-Sinai entropy [3], and this implies that (A^G, μ) has completely positive entropy and hence is uniformly mixing, by the result of [11].

Thus if G is amenable then completely positive entropy and uniform mixing are both equivalent to the existence of a sofic approximation and a locally and empirically convergent sequence of measures over it which is uniformly model-mixing. If these conditions hold, then we expect that one can actually find a locally and empirically convergent and uniformly model-mixing sequence of measures over *any* sofic approximation to G. This should follow using a similar kind of decomposition of the sofic approximants into Følner sets as in Bowen's proof in [3]. However, we have not explored this argument in detail.

Definition 3.1 applies to a shift-system with a finite state space. It can be transferred to an abstract measure-preserving G-action on (X,μ) by fixing a choice of finite measurable partition of X. However, in order to study actions which do not admit a finite generating partition, it might be worth looking for an extension of Definition 3.1 to G-processes with arbitrary compact metric state spaces, similarly to the setting in [1]. We also do not pursue this generalization here.

4 Proof of Theorem 1.2

We will use basic facts about the Shannon entropy of observables (i.e. random variables with finite range), for which we refer the reader to Chapter 2 of [4]. Let $\Sigma = (\sigma_n : G \to \mathrm{Sym}(V_n))$, (A^G,μ) and $(\mu_n)_{n=1}^\infty$ be as in the statement of Theorem 1.2. The following is the 'finitary' model-measure analog of Lemma 5.1 in [5].

Lemma 4.1. *Let $F \subseteq G$ be finite. Let B be a finite set and let $\phi : A^G \to B$ be an F-local observable. Let $S_n \subseteq V_n$ be a sequence of sets such that $|S_n| \gtrsim |V_n|$. Then we have*

$$\mathrm{H}(\mu_F) - \frac{1}{|S_n|}\mathrm{H}(\pi_{\sigma_n^F(S_n)*}\mu_n) \geq \mathrm{H}_\mu(\phi) - \frac{1}{|S_n|}\mathrm{H}\big(\pi_{S_n*}\phi_*^{\sigma_n}\mu_n\big) - o(1).$$

Proof of Lemma 4.1. Let $\theta : A^F \to B$ be a map with $\theta \circ \pi_F = \phi$. Fix $n \in \mathbb{N}$ and $S \subseteq V_n$. Let $\alpha = \pi_{\sigma_n^F(S)} : A^{V_n} \to A^{\sigma_n^F(S)}$ and let $\beta = \pi_S \circ \phi^{\sigma_n} : A^{V_n} \to B^S$. For $s \in S$ let $\alpha_s = \Pi^{\sigma_n}_{s,F} : A^{V_n} \to A^F$ and let $\beta_s = \theta \circ \Pi^{\sigma_n}_{s,F} : A^{V_n} \to B$. Then we have $\alpha = (\alpha_s)_{s\in S}$ and $\beta = (\beta_s)_{s\in S}$. Enumerate $S = (s_k)_{k=1}^m$ and write $\alpha_{s_k} = \alpha_k$. All entropies in the following display are computed with respect to μ_n. We have

$$\begin{aligned}
\mathrm{H}(\alpha) &= \mathrm{H}(\alpha_1,\ldots,\alpha_m)\\
&= \mathrm{H}(\alpha_1) + \sum_{k=1}^{m-1}\mathrm{H}(\alpha_{k+1}|\alpha_1,\ldots,\alpha_k)\\
&= \mathrm{H}(\alpha_1,\beta_1) + \sum_{k=1}^{m-1}\mathrm{H}(\alpha_{k+1},\beta_{k+1}|\alpha_1,\ldots,\alpha_k)\\
&= \mathrm{H}(\beta_1) + \mathrm{H}(\alpha_1|\beta_1) + \sum_{k=1}^{m-1}\mathrm{H}(\beta_{k+1}|\alpha_1,\ldots,\alpha_k) + \sum_{k=1}^{m-1}\mathrm{H}(\alpha_{k+1}|\beta_{k+1},\alpha_1,\ldots,\alpha_k)\\
&\leq \mathrm{H}(\beta_1) + \sum_{k=1}^{m-1}\mathrm{H}(\beta_{k+1}|\beta_1,\ldots,\beta_k) + \sum_{k=1}^{m}\mathrm{H}(\alpha_k|\beta_k)\\
&= \mathrm{H}(\beta) + \sum_{k=1}^{m}\mathrm{H}(\alpha_k|\beta_k).
\end{aligned}$$

Let ι be the identity map on A^F. Then

$$\begin{aligned}
|S| \cdot \mathrm{H}(\mu_F) - \mathrm{H}(\pi_{\sigma_n^F(S)*}\mu_n) &= |S| \cdot \mathrm{H}_{\mu_F}(\iota) - \mathrm{H}_{\mu_n}(\alpha) \\
&\geq |S| \cdot \mathrm{H}_{\mu_F}(\theta) + |S| \cdot \mathrm{H}_{\mu_F}(\iota|\theta) - \mathrm{H}_{\mu_n}(\beta) - \sum_{s\in S} \mathrm{H}_{\mu_n}(\alpha_s|\beta_s) \\
&= |S| \cdot \mathrm{H}_\mu(\phi) - \mathrm{H}\big(\pi_{S*}\phi^{\sigma_n}_* \mu_n\big) + |S| \cdot \mathrm{H}_{\mu_F}(\iota|\theta) - \sum_{s\in S} \mathrm{H}_{\mu_n}(\alpha_s|\beta_s). \qquad (4.1)
\end{aligned}$$

Now allowing n to vary, let $S_n \subseteq V_n$ be a sequence of sets such that $|S_n| \gtrsim |V_n|$. Write $\nu_n = \pi_{\sigma_n^F(S_n)*}\mu_n$. Let $s \in S_n$ be such that the obvious map from F to $\sigma_n^F(s)$ is injective. Then the function $\overline{a} \mapsto \Pi^{\sigma_n}_{s,F}(\overline{a})$ provides an identification of $A^{\sigma_n^F(s)}$ with A^F. This identification sends α_s to ι and β_s to θ. When n is large the $\sigma_n^F(s)$ marginal of μ_n will resemble μ_F for most $s \in S_n$. Since α_s and β_s are $\pi_{\sigma_n^F(s)}$ measurable this implies that $\mathrm{H}_{\mu_F}(\iota|\theta) \approx \mathrm{H}_{\nu_n}(\alpha_s|\beta_s)$ for most s. More precisely, we can find a sequence of sets $C_n \subseteq S_n$ with

$$|C_n| = (1 - o(1))|S_n|$$

such that

$$\max_{s\in C_n} \big|\mathrm{H}_{\mu_F}(\iota|\theta) - \mathrm{H}_{\nu_n}(\alpha_s|\beta_s)\big| = o(1).$$

Thus

$$\begin{aligned}
\left| |S_n| \cdot \mathrm{H}_{\mu_F}(\iota|\theta) - \sum_{s\in S_n} \mathrm{H}_{\nu_n}(\alpha_s|\beta_s) \right| &\leq \sum_{s\in C_n} \big|\mathrm{H}_{\mu_F}(\iota|\theta) - \mathrm{H}_{\nu_n}(\alpha_s|\beta_s)\big| + \sum_{s\in S_n\setminus C_n} \big|\mathrm{H}_{\mu_F}(\iota|\theta) - \mathrm{H}_{\nu_n}(\alpha_s|\beta_s)\big| \\
&= o(|S_n|).
\end{aligned}$$

The lemma then follows from (4.1) and the above. $\square$

Recall that for a measure space (X, μ) and two observables α and β on X the Rokhlin distance between α and β is defined by

$$d^{\mathrm{Rok}}_\mu(\alpha, \beta) = \mathrm{H}_\mu(\alpha|\beta) + \mathrm{H}_\mu(\beta|\alpha).$$

This is a pseudometric on the space of observables on X. An easy computation shows that if $\alpha_1, \dots, \alpha_n$ and $\beta_1, \dots, \beta_n$ are two families of observables on X then

$$d^{\mathrm{Rok}}_\mu((\alpha_1, \dots, \alpha_n), (\beta_1, \dots, \beta_n)) \leq \sum_{k=1}^n d^{\mathrm{Rok}}_\mu(\alpha_k, \beta_k).$$

Lemma 4.2. *Let $\phi, \psi : A^G \to B$ be two local observables. Let $S_n \subseteq V_n$ be a sequence of sets with $|S_n| \gtrsim |V_n|$. Then we have*

$$\frac{1}{|S_n|}\big|\mathrm{H}(\pi_{S_n*}\phi^{\sigma_n}_*\mu_n) - \mathrm{H}(\pi_{S_n*}\psi^{\sigma_n}_*\mu_n)\big| \leq d^{\mathrm{Rok}}_\mu(\phi, \psi) + o(1).$$

Proof. Let $\alpha_n = \pi_{S_n} \circ \phi^{\sigma_n} : A^{V_n} \to B^{S_n}$ and let $\beta_n = \pi_{S_n} \circ \psi^{\sigma_n} : A^{V_n} \to B^{S_n}$. Let F be a finite subset of G such that both ϕ and ψ are F-local. Let $\theta : A^F \to B$ be a map such that $\theta \circ \pi_F = \phi$ and let $\kappa : A^F \to B$

be a map such that $\kappa \circ \pi_F = \psi$. For $s \in S_n$ let $\alpha_{n,s} = \theta \circ \Pi^{\sigma_n}_{s,F} : A^{V_n} \to B$ so that $\alpha_n = (\alpha_{n,s})_{s \in S_n}$. Also let $\beta_{n,s} = \kappa \circ \Pi^{\sigma_n}_{s,F} : A^{V_n} \to B$. Then we have

$$\begin{aligned}
\frac{1}{|S_n|}\left|\mathrm{H}(\pi_{S_n *}\phi^{\sigma_n}_* \mu_n) - \mathrm{H}(\pi_{S_n *}\psi^{\sigma_n}_* \mu_n)\right| &= \frac{1}{|S_n|}\left|\mathrm{H}_{\mu_n}(\alpha_n) - \mathrm{H}_{\mu_n}(\beta_n)\right| \\
&\leq \frac{1}{|S_n|} \cdot d^{\mathrm{Rok}}_{\mu_n}(\alpha_n, \beta_n) \\
&= \frac{1}{|S_n|} \cdot d^{\mathrm{Rok}}_{\mu_n}\big((\alpha_{n,s})_{s\in S_n}, (\beta_{n,s})_{s \in S_n}\big) \\
&\leq \frac{1}{|S_n|} \sum_{s \in S_n} d^{\mathrm{Rok}}_{\mu_n}(\alpha_{n,s}, \beta_{n,s}) \qquad (4.2)
\end{aligned}$$

If the map $g \mapsto \sigma^g_n \cdot s$ is injective on F, we can identify $A^{\sigma^F_n(s)}$ with A^F and thereby identify $\alpha_{n,s}$ with θ and $\beta_{n,s}$ with κ. Note that

$$d^{\mathrm{Rok}}_{\mu_F}(\theta, \kappa) = d^{\mathrm{Rok}}_{\mu}(\phi, \psi).$$

It follows that for any $\epsilon > 0$ we can find a weak star neighborhood $\mathcal{O}$ of μ such that if $s \in S_n$ is such that $(\Pi^{\sigma_n}_s)_* \mu_n \in \mathcal{O}$ then

$$\left| d^{\mathrm{Rok}}_{\mu_n}(\alpha_{n,s}, \beta_{n,s}) - d^{\mathrm{Rok}}_{\mu}(\phi, \psi) \right| < \epsilon.$$

Thus, since μ_n locally and empirically converges to μ, there are sets $C_n \subseteq S_n$ with $|C_n| = (1 - o(1))|S_n|$ such that

$$\max_{s \in C_n} \left| d^{\mathrm{Rok}}_{\mu_n}(\alpha_{n,s}, \beta_{n,s}) - d^{\mathrm{Rok}}_{\mu}(\phi, \psi) \right| = o(1). \qquad (4.3)$$

The lemma now follows from (4.2) and (4.3). □

Corollary 4.1. *Let $\left(\phi_m : A^G \to B\right)_{m=1}^{\infty}$ be a sequence of local observables and let $\phi : A^G \to B$ be a local observable. Let $S_n \subseteq V_n$ be a sequence of sets with $|S_n| \gtrsim |V_n|$. Then if $(m_n)_{n=1}^{\infty}$ increases to infinity at a slow enough rate we have*

$$\frac{1}{|S_n|}\left|\mathrm{H}(\pi_{S_n *}\phi^{\sigma_n}_* \mu_n) - \mathrm{H}(\pi_{S_n *}\phi^{\sigma_n}_{m_n *}\mu_n)\right| \leq d^{\mathrm{Rok}}_{\mu}(\phi, \phi_{m_n}) + o(1).$$

Proof of Theorem 1.2. Let B be a finite set and let $\psi : A^G \to B$ be an observable with $\mathrm{H}_\mu(\psi) > 0$. Let $(\phi_m)_{m=1}^{\infty}$ be an AL approximating sequence for ψ rel μ (see Definition 4.4 in [1]). Then the sequence ϕ_m converges to ψ in d^{Rok}_{μ}. In particular, ϕ_m is a Cauchy sequence and so we can find $M \in \mathbb{N}$ so that for all $m \geq M$ we have

$$d^{\mathrm{Rok}}_{\mu}(\phi_m, \phi_M) \leq \frac{\mathrm{H}_\mu(\psi)}{8}. \qquad (4.4)$$

We will also assume M is large enough that

$$\mathrm{H}_\mu(\phi_M) \geq \frac{\mathrm{H}_\mu(\psi)}{2}. \qquad (4.5)$$

Let F be a finite subset of G such that ϕ_M is F-local. Then Definition 3.1 provides an $r < \infty$ and a sequence of subsets $W_n \subseteq V_n$ such that $|W_n| = (1 - o(1))|V_n|$ and if $S \subseteq W_n$ is r-separated then

$$\mathrm{H}(\mu_F) - \frac{1}{|S|}\mathrm{H}(\pi_{\sigma^F_n(S)*}\mu_n) \leq \frac{\mathrm{H}_\mu(\phi_M)}{2}. \qquad (4.6)$$

Let $K = |B_\rho(1_G, r)|$. Since σ_n is a sofic approximation there are sets $W_n' \subseteq V_n$ with $|W_n'| = (1 - o(1))|V_n|$ such that if $w \in W_n'$ then the ρ_n ball of radius r around w has cardinality at most K. Write $Y_n = W_n \cap W_n'$ and note that we have $|Y_n| = (1 - o(1))|V_n|$. For each n let S_n be an r-separated subset of Y_n with maximal cardinality. Then $Y_n \subseteq \bigcup_{s \in S_n} B_{\rho_n}(s, r)$ so that

$$|S_n| \geq \frac{|Y_n|}{K} = (1 - o(1))\frac{|V_n|}{K}. \tag{4.7}$$

By Lemma 4.1 and (4.6) we have

$$\mathrm{H}_\mu(\phi_M) - \frac{1}{|S_n|}\mathrm{H}\big(\pi_{S_n *}\phi^{\sigma_n}_{M*}\mu_n\big) - o(1) \leq \frac{\mathrm{H}_\mu(\phi_M)}{2}$$

so that from (4.5) we have

$$\frac{\mathrm{H}_\mu(\psi)}{4} - o(1) \leq \frac{1}{|S_n|}\mathrm{H}\big(\pi_{S_n *}\phi^{\sigma_n}_{M*}\mu_n\big). \tag{4.8}$$

By Proposition 5.15 in [1] if $(m_n)_{n=1}^\infty$ increases to infinity at a slow enough rate then $(\phi^{\sigma_n}_{m_n})_*\mu_n$ will locally and empirically converge to $\psi^G_*\mu$. Since A is finite, by the same argument as for Proposition 8.1 in [1] we have

$$\begin{aligned}\underline{\mathrm{h}}^{\mathrm{q}}_\Sigma\big(\psi^G_*\mu\big) &\geq \sup_{\epsilon>0}\liminf_{n\to\infty}\frac{1}{|V_n|}\log\operatorname{cov}_\epsilon\big((\phi^{\sigma_n}_{m_n})_*\mu_n\big)\\ &\geq \liminf_{n\to\infty}\frac{1}{|V_n|}\,\mathrm{H}\big((\phi^{\sigma_n}_{m_n})_*\mu_n\big)\end{aligned} \tag{4.9}$$

where the second inequality follows from Lemma 2.1. We also assume that $(m_n)_{n=1}^\infty$ increases slowly enough for Corollary 4.1 to hold. By (4.4) we have

$$\left|\frac{1}{|S_n|}\mathrm{H}\big(\pi_{S_n *}\phi^{\sigma_n}_{M*}\mu_n\big) - \frac{1}{|S_n|}\mathrm{H}\big(\pi_{S_n *}(\phi^{\sigma_n}_{m_n})_*\mu_n\big)\right| \leq \frac{\mathrm{H}_\mu(\psi)}{8} + o(1).$$

Combining this with (4.8) we see that

$$\frac{1}{|S_n|}\mathrm{H}\big(\pi_{S_n *}(\phi^{\sigma_n}_{m_n})_*\mu_n\big) \geq \frac{\mathrm{H}_\mu(\psi)}{8} - o(1).$$

By the above and (4.7) we have that for all sufficiently large n,

$$\mathrm{H}\big((\phi^{\sigma_n}_{m_n})_*\mu_n\big) \geq \frac{\mathrm{H}_\mu(\psi)}{8K+1}|V_n| \tag{4.10}$$

Theorem 1.2 now follows from (4.9) and (4.10). □

5 Proof of Theorem 1.3

Let $(A^{\mathbb{Z}}, \nu)$ be a uniformly mixing $\mathbb{Z}$-process, and for each positive integer l let ν_l be the marginal of ν on A^l. Let $\Sigma = (\sigma_n : G \to \mathrm{Sym}(V_n))$ be an arbitrary sofic approximation to G. Let $h \in G$ have infinite order and write $H = \langle h \rangle \cong \mathbb{Z}$. We construct a measure μ_n on A^{V_n} for each $n \in \mathbb{N}$. We will later show that the sequence $(\mu_n)_{n=1}^{\infty}$ is uniformly model-mixing and locally and empirically converges to μ over Σ.

We first construct a measure μ_n^l on A^{V_n} for each pair (n, l) with l much smaller than n. For a given n, the single permutation σ_n^h partitions V_n into a disjoint union of cycles. Since h has infinite order and Σ is a sofic approximation, once n is large most points will be in very long cycles. In particular we assume that most points are in cycles with length much larger than l. Partition the cycles into disjoint paths so that as many of the paths have length l as possible, and let $\mathcal{P}_n^l = (P_{n,1}^l, \dots, P_{n,k_n}^l)$ be the collection of all length-l paths that result (so $\mathcal{P}_n^l$ is not a partition of the whole of V_n, but covers most of it). Fix any element $\overline{a}_0 \in A^{V_n}$ and define a random element $\overline{a} \in A^{V_n}$ by choosing each restriction $\overline{a} \restriction_{P_{n,i}^l}$ independently with the distribution of ν_l and extending to the rest of V_n according to $\overline{a}_0$. Let μ_n^l be the law of this $\overline{a}$.

Now let $(l_n)_{n=1}^{\infty}$ increase to infinity at a slow enough rate that the following two conditions are satisfied:

(a) The number of points of V_n that lie in some member of the family $\mathcal{P}_n^{l_n}$ is $(1 - o(1))|V_n|$.

(b) Whenever $g, g' \in G$ lie in distinct right cosets of H, so that $g^{-1}h^p g' \neq 1_G$ for all $p \in \mathbb{Z}$, we have

$$|\{v \in V_n : (\sigma_n^g)^{-1}(\sigma_n^h)^p \sigma_n^{g'} \cdot v = v \text{ for some } p \in \{-l_n, \dots, l_n\}\}| = o(|V_n|)$$

Set $\mu_n = \mu_n^{l_n}$. We separate the proof that $(\mu_n)_{n=1}^{\infty}$ has the required properties into two lemmas.

Lemma 5.1. *$(\mu_n)_{n=1}^{\infty}$ locally and empirically converges to μ over Σ.*

Proof of Lemma 5.1. Since (A^G, μ) is ergodic, by Corollary 5.6 in [1] it suffices to show that μ_n locally weak star converges to μ. For a set $I \subseteq \mathbb{Z}$ write $h^I = \{h^i : i \in I\}$. Fix a finite set $F \subseteq G$. By enlarging F if necessary we can assume there is an interval $I \subseteq \mathbb{Z}$ such that $F = \bigcup_{k=1}^{m} h^I t_k$ for $t_1, \dots, t_m$ in some transversal for the right cosets of H in G. For each $g \in F$ let j_g be a fixed element of A. Let $B \subseteq A^G$ be defined by

$$B = \left\{a \in A^G : a(g) = j_g \text{ for all } g \in F\right\}$$

and let $\epsilon > 0$. Then sets such as

$$\mathcal{O} = \left\{\eta \in \mathrm{Prob}(A^G) : \eta(B) \approx_\epsilon \mu(B)\right\}$$

form a subbasis of neighborhoods around μ. It therefore suffices to show that when n is large we have $(\Pi_v^{\sigma_n})_* \mu_n \in \mathcal{O}$ with high probability in the choice of $v \in V_n$.

For $k \in \{1, \dots, m\}$ let
$$B_k = \left\{x \in A^{\mathbb{Z}} : x(i) = j_{h^i t_k} \text{ for all } i \in I\right\}.$$

Note that μ is defined in such a way that $\mu(B) = \prod_{i=1}^{k} \nu(B_k)$. Now, let W_n be the set of all points $v \in V_n$ such that the following conditions hold.

(i) The map $g \mapsto \sigma_n^g \cdot v$ is injective on F.

(ii) $\sigma_n^{h^i t_k} \cdot v = (\sigma_n^h)^i \sigma_n^{t_k} \cdot v$ for all $i \in I$ and $k \in \{1, \ldots, m\}$.

(iii) For all pairs $g, g' \in F$, $\sigma_n^g \cdot v$ is in the same path as $\sigma_n^{g'} \cdot v$ if and only if g and g' lie in the same right coset of H. In particular, each of the images $\sigma_n^g \cdot v$ for $g \in F$ is contained in some member of $\mathcal{P}_n^{l_n}$.

We claim that $|W_n| = (1 - o(1))|V_n|$. Clearly Conditions (i) and (ii) are satisfied with high probability in v, and so is the last part of Condition (iii), by Condition (a) in the choice of $(l_n)_{n=1}^\infty$.

Fix $g, g' \in F$ and suppose that g and g' are in the same coset of H, so that we have $g = h^i t_k$ and $g' = h^{i'} t_k$ for some $k \in \{1, \ldots, m\}$ and $i, i' \in I$. If v satisfies Condition (ii) then we have

$$(\sigma_n^h)^{i'-i} \sigma_n^g \cdot v = (\sigma_n^h)^{i'-i} (\sigma_n^h)^i \sigma_n^{t_k} \cdot v = (\sigma_n^h)^{i'} \sigma_n^{t_k} \cdot v = \sigma_n^{g'} \cdot v$$

so that $\sigma_n^g \cdot v$ and $\sigma_n^{g'} \cdot v$ will lie in the same path assuming that $\sigma_n^{t_k} \cdot v$ is not one of the first or last $|I|$ elements of its path. Note that for any $v \in V_n$ we have

$$\left|\left\{w : \sigma_n^{t_k} \cdot w = v \text{ for some } k \in \{1, \ldots, m\}\right\}\right| \leq m.$$

It follows that the number of points $v \in V_n$ such that $\sigma_n^{t_k} \cdot v$ is one of the first or last $|I|$ elements of a path is at most $2mp_n|I| + o(|V_n|)$ where p_n is the total number of paths in V_n. By Condition (a) in the choice of $(l_n)_{n=1}^\infty$, most of V_n is covered by paths whose lengths increase to infinity. Since also $p_n = o(V_n)$, it follows that $\sigma_n^g \cdot v$ lies in the same path as $\sigma_n^{g'} \cdot v$ with high probability in v.

On the other hand, suppose that g and g' are in distinct cosets of H. Assume that $\sigma_n^g \cdot v$ and $\sigma_n^{g'} \cdot v$ are in the same path. Then there is $p \in \{-l_n, \ldots, l_n\}$ with $\sigma_n^g \cdot v = (\sigma_n^h)^p \sigma_n^{g'} \cdot v$, and hence $(\sigma_n^g)^{-1} (\sigma_n^h)^p \sigma_n^{g'} \cdot v = v$. By Condition (b) in the choice of $(l_n)_{n=1}^\infty$ there are only $o(|V_n|)$ vertices v for which this holds. Thus we have established the claim.

Now let $v \in W_n$. We have

$$(\Pi_v^{\sigma_n})_* \mu_n(B) = \mu_n\left(\left\{\overline{a} \in A^{V_n} : \overline{a}(\sigma_n^g \cdot v) = j_g \text{ for all } g \in F\right\}\right).$$

For each $k \in \{1, \ldots, m\}$ the set $\{(\sigma_n^h)^i \sigma_n^{t_k} \cdot v : i \in I\}$ is contained in a single path. Since the marginal of μ_n on each path is ν_{l_n} the probability that

$$\overline{a}\left((\sigma_n^h)^i \sigma_n^{t_k} \cdot v\right) = j_{h^i t_k}$$

for all $i \in I$ is equal to $\nu_{l_n}(B_k) = \nu(B_k)$. On the other hand, the marginals of μ_n on distinct paths are independent, so the probability that $\overline{a}(\sigma_n^g \cdot v) = j_g$ for all $g \in F$ is actually equal to $\prod_{i=1}^{k} \nu(B_k)$. $\square$

If $(A^{\mathbb{Z}}, \nu)$ is weakly mixing, then so is the co-induced G-action. In particular, this certainly holds if $(A^{\mathbb{Z}}, \nu)$ is uniformly mixing. Therefore we may immediately promote Lemma 5.1 to the fact that $(\mu_n)_{n=1}^\infty$ locally and doubly empirically converges to μ over Σ, by Lemma 5.15 of [1]. In fact, we suspect that local and double empirical convergence holds here whenever $(A^{\mathbb{Z}}, \nu)$ is ergodic.

Lemma 5.2. *$(\mu_n)_{n=1}^\infty$ is uniformly model-mixing.*

Proof of Lemma 5.2. Let $F \subseteq G$ be finite and let $\epsilon > 0$. Again decompose $F = \bigcup_{k=1}^{m} h^I t_k$ for some interval $I \subseteq \mathbb{Z}$ and elements $t_k \in T$. Note that the restriction of the metric ρ to H is a proper right invariant metric on $H \cong \mathbb{Z}$, even though it might be different from the usual metric on $\mathbb{Z}$. Thus since ν is uniformly mixing we can find some $r_0 < \infty$ such that if $(I_j)_{j=1}^q$ is a family of intervals in $\mathbb{Z}$ which are each of length $|I|$ and are pairwise at distance at least r_0 then writing $K = \bigcup_{j=1}^{q} I_j$ we have

$$\mathrm{H}(\nu_K) \geq q \cdot \left(\mathrm{H}(\nu_I) - \frac{\epsilon}{m}\right). \tag{5.1}$$

Let $r < \infty$ be large enough that for all $g, g' \in G$ if $\rho(g, g') \geq r$ then $\rho(fg, f'g') \geq r_0$ for all $f, f' \in F$. Such a choice of r is possible since by right-invariance of ρ we have $\rho(fg, g) = \rho(f, 1_G)$ and $\rho(f'g', g') = \rho(f', 1_G)$. Let W_n be as in the proof of Lemma 5.1 and recall that $|W_n| = (1 - o(1))|V_n|$. Let $S \subseteq W_n$ be r-separated according to ρ_n.

Fix a path $P \in \mathcal{P}_n^{l_n}$ and let S_P be the set of points $v \in S$ such that $\sigma_n^{t_{k(v)}} \cdot v \in P$ for some $k(v) \in \{1, \ldots, m\}$. Since $S \subseteq W_n$, Condition (iii) from the previous proof implies that

$$\sigma_n^F(S) \cap P = \bigcup_{v \in S_P} \{(\sigma_n^h)^i \sigma_n^{t_{k(v)}} \cdot v : i \in I\}.$$

Each of the sets in the latter union is an interval of length $|I|$ in P and by our choice of r these are pairwise at distance r_0 in ρ_n restricted to P. Since the marginal of μ_n on P is equal to ν_{n_l}, (5.1) implies that

$$\mathrm{H}(\pi_{(\sigma_n^F(S) \cap P)*}\mu_n) \geq |S_P| \cdot \left(\mathrm{H}(\nu_I) - \frac{\epsilon}{m}\right).$$

Since the marginals of μ_n on distinct paths are independent, this implies that

$$\mathrm{H}(\pi_{\sigma_n^F(S)*}\mu_n) \geq \left(\sum_{P \in \mathcal{P}_n^{l_n}} |S_P|\right) \cdot \left(\mathrm{H}(\nu_I) - \frac{\epsilon}{m}\right). \tag{5.2}$$

By Condition (iii) in the definition of W_n, each $v \in S$ appears in S_P for exactly m paths P. Therefore

$$\sum_{P \in \mathcal{P}_n^{l_n}} |S_P| = m \cdot |S|. \tag{5.3}$$

Now $\mathrm{H}(\mu_F) = m \cdot \mathrm{H}(\nu_I)$ so from (5.2) and (5.3) we have

$$\mathrm{H}(\pi_{\sigma_n^F(S)*}\mu_n) \geq |S| \cdot (\mathrm{H}(\mu_F) - \epsilon)$$

as required. □

Proof of Theorem 1.3. Theorem 1.3 now follows from Theorem 1.2 and Lemmas 5.1 and 5.2. □

6 Proof of Theorem 1.1

Proof of Theorem 1.1. This part of the argument is essentially the same as the corresponding part of [5]. Consider the family of uniformly mixing $\mathbb{Z}$-processes $\{(4^{\mathbb{Z}}, \nu_\omega) : \omega \in 2^{\mathbb{N}}\}$ constructed in Section 6 of [5]. Fix an isomorphic copy H of $\mathbb{Z}$ in G and let $\mu_\omega = \mathrm{CInd}_H^G(\nu_\omega)$. By Theorems 1.2 and 1.3 the process $(4^G, \mu_\omega)$ has completely positive model-measure sofic entropy. Note that the restriction of the G-action to H is a permuted power of the original $\mathbb{Z}$-process in the sense of Definition 6.5 from [5]. Thus by Proposition 6.6 in that reference, the processes $\{(4^G, \mu_\omega) : \omega \in 2^{\mathbb{N}}\}$ are pairwise nonisomorphic.

Suppose toward a contradiction that for some ω, $(4^G, \mu_\omega)$ is a factor of a Bernoulli shift (Z^G, η^G) over some standard probability space (Z, η). Let $\psi : Z^G \to 4^G$ be an equivariant measurable map with $\psi_*\eta^G = \mu_\omega$. Note that the restricted right-shift action $H \curvearrowright (Z^G, \eta^G)$ is still isomorphic to a Bernoulli shift and ψ is still a factor map from this process to the restricted action $H \curvearrowright (4^G, \mu_\omega)$. Thus the latter $\mathbb{Z}$-process is isomorphic to a Bernoulli shift and so is its factor $(4^{\mathbb{Z}}, \nu_\omega)$. This contradicts Corollary 6.4 in [5]. $\square$

References

[1] T. Austin. Additivity properties of sofic entropy and measures on model spaces. preprint, http://arxiv.org/abs/1510.02392, 2015.

[2] L. Bowen. Measure conjugacy invariants for actions of countable sofic groups. *J. Amer. Math. Soc.*, pages 217–245, 2010.

[3] L. Bowen. Sofic entropy and amenable groups. *Ergodic Theory and Dynamical Systems*, 32(2):427–466, 2012.

[4] Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, second edition, 2006.

[5] A. Dooley, V. Golodets, D. Rudolph, and S. Sinel'shchikov. Non-Bernoulli systems with completely positive entropy. *Ergodic Theory and Dynamical Systems*, 28:87–124, 2007.

[6] Valentin Ya. Golodets and Sergey D. Sinel'shchikov. Complete positivity of entropy and non-Bernoullicity for transformation groups. *Colloq. Math.*, 84/85(part 2):421–429, 2000. Dedicated to the memory of Anzelm Iwanik.

[7] A.S. Kechris. *Global aspects of ergodic group actions*, volume 160 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2010.

[8] D. Kerr. Sofic measure entropy via finite partitions. *Groups Geom. Dyn.*, 7:617–632, 2013.

[9] D. Kerr and H. Li. Entropy and the variational principle for actions of sofic groups. *Inventiones Mathematicae*, 186:501–558, 2011.

[10] David Kerr. Bernoulli actions of sofic groups have completely positive entropy. *Israel J. Math.*, 202(1):461–474, 2014.

[11] D. Rudolph and B. Weiss. Entropy and mixing for amenable group actions. *Annals of Mathematics*, 151(3):1119–1150, 2000.

[12] B. Seward. Krieger's finite generator theorem for ergodic actions of countable groups I. preprint, http://arxiv.org/abs/1405.3604, 2014.

[13] B. Seward. Krieger's finite generator theorem for ergodic actions of countable groups II. preprint, http://arxiv.org/abs/1501.03367, 2014.

[14] Benjamin Weiss. Actions of amenable groups. In *Topics in dynamics and ergodic theory*, volume 310 of *London Math. Soc. Lecture Note Ser.*, pages 226–262. Cambridge Univ. Press, Cambridge, 2003.

Courant Institute of Mathematical Sciences
New York University
New York NY, 10012
`tim@cims.nyu.edu`

Department of Mathematics
California Institute of Technology
Pasadena CA, 91125
`pjburton@caltech.edu`